UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK
---x

THE NEW YORK TIMES COMPANY,   :   23-cv-11195 (SHS) (OTW)
                                    :   24-cv-3285 (SHS) (OTW)

         Plaintiff,   :

                                      :   **OPINION & ORDER**

         -against-   :

MICROSOFT CORPORATION, et al.,   :

         Defendants.   :

                                      :

                                      :
---x

        ONA T. WANG, **United States Magistrate Judge**:

        Pending now before the Court is Defendant OpenAI's request for a protective order

prohibiting the disclosure of certain "highly confidential documents" to Dr. Ricardo Baeza-Yates

("Dr. Baeza-Yates"), an expert retained by Plaintiffs The New York Times and Daily News

(collectively, "Plaintiffs").[1] For the reasons discussed below, OpenAI's motion is **DENIED.**

 I.   **BACKGROUND**

        The Court assumes familiarity with the facts of this case. On October 18, 2024, OpenAI

filed a motion seeking a protective order to prevent the disclosure of confidential documents to

Dr. Baeza-Yates, the co-founder and Chief Scientific Officer of Theodora.AI. (24-CV-3285,

---

[1] While the New York Times and Daily News cases were consolidated at the September 12, 2024, status conference, (*see* ECF 243), OpenAI's motion for a protective order was filed only on the Daily News docket. (24-CV-3285, ECF 163). Plaintiffs filed a responsive letter on both dockets. (ECF 271; 24-CV-3285, ECF 172). Plaintiffs then filed a letter on the New York Times docket informing the Court that the parties met and conferred by remained at an impasse. (ECF 469).

ECF 163).[2] Plaintiffs filed their opposition on October 22, 2024, arguing that: (1) there is no risk

of competitive harm to OpenAI from disclosing documents in this case to Dr. Baeza-Yates; and

(2) the *Flores* balancing test, if applicable, favors rejecting the requested protective order.

(ECF 271; 24-CV-3285, 172). On March 3, 2025, Plaintiffs filed an additional letter seeking

resolution of the dispute. (ECF 469). The Court notes that that the parties did not include this

issue in their last two joint dispute charts filed on January 17, 2025, and February 13, 2025.

(*See* ECF 431, 462).

II.   ANALYSIS

OpenAI argues that a protective order is needed because Theodora.AI is a competitor

that "works with revolutionary AI to create deep tech and develop AI-based tools," and implies

that the disclosure of such documents to Theodora.AI's CSO may cause significant competitive

harm to OpenAI. (24-CV-3285, ECF 163, internal citations omitted).

Under Fed. R. Civ. P. 26(c)(1)(G), a court may issue a protective order to require "a trade

secret or other confidential research, development, or commercial information not be revealed

or be revealed only in a specific way." The party seeking a protective order has the burden of

establishing "good cause" for the order, which is established when there is a "particular need

for protection" and when a party can show that "disclosure will result in a clearly defined,

specific and serious injury." Ultimately, though, the court has to weigh the interests of both

sides. *Flores v. Stanford*, 18-CV-02468, 2021 WL 4441614, at *4 (S.D.N.Y. Sept. 28, 2021). If the

commercially sensitive materials are relevant to a party's claim, the court must balance the

---

[2] Though the motion refers to Dr. Baeza-Yates's as Daily News's expert, The New York Times subsequently
disclosed Dr. Baeza-Yates as an expert as well.

"[receiving] party's need to have its chosen expert review the materials against the producing

party's interest in protecting the materials from a potential competitor." *Id.* at *4.[3]

Courts weighing whether to enter a protective order for commercially sensitive

materials consider following factors: (1) whether the person receiving the confidential

information is involved in competitive decision making or scientific research relating to the

subject matter of the materials; (2) the risk of inadvertent disclosure of proprietary

information; (3) the hardship imposed by the restriction; (4) the timing of the remedy; and (5)

the scope of the remedy. *Flores*, 2021 WL 4441614, at *4; *see also Uniroyal Chem. Co. Inc. v.*

Syngenta Crop Prot., 224 F.R.D. 53, 57 (D. Conn. 2004). Court must also consider the "specific

expertise" of the consultant and whether other consultants "possess similar expertise." *Flores*,

2021 WL 4441614, at *4.

As an initial matter, OpenAI's motion should be denied because OpenAI failed to

articulate any concrete injury that may result from disclosure of documents to Dr. Baeza-Yates.

OpenAI only offers conclusory statements that disclosing documents to Dr. Baeza-Yates will

result in commercial harm and that the risk of harm outweighs Plaintiffs' need to disclose, but

---

[3] The existing Protective Order in these consolidated cases also sets forth:

> [When a party seeks] to disclose to a testifying or consulting expert any
> information or item that has been designated 'HIGHLY CONFIDENTIAL –
> ATTORNEYS' EYES ONLY' or 'HIGHLY CONFIDENTIAL – SOURCE CODE' [then the]
> party opposing disclosure to the expert shall bear the burden of proving that the
> risk of harm that the disclosure would entail (under the safeguards proposed)
> outweighs the recipient's need to disclose the Protected Discovery Material to
> its expert.

(ECF 127 ¶ 17(b); 24-cv-3285, ECF 129 ¶ 16(b)).

The parties' original protective order in this case was filed on May 31, 2024. (ECF 127). The parties recently
submitted a revised protective order, which the Court entered on March 18, 2025. (ECF Nos. 473, 474). The revised
protective order does not affect the analysis in this order.

OpenAI does not define or quantify what harm is likely to result. (24-CV-3285, ECF 163 at 1).

Notwithstanding, the Court will analyze each of the *Flores* factors in turn.

A. Competitive Risk

Here, disclosing confidential documents to Dr. Baeza-Yates is unlikely to result in

competitive harm to OpenAI simply because Theodora.AI is not a meaningful competitor.

Unlike OpenAI, which was valued at $90 billion and generated revenue to the tune of $1 billion

just two years ago with innumerable AI-related products, only a portion of which are at issue in

the current cases, (e.g., GPT models), Theodora.AI is a small, six-person Chilean company that

only works on fine-tuning others' LLMs to mitigate Spanish language bias. (ECF Nos. 1 ¶ 57, 271

at 1). OpenAI does not allege to have even a single product that is aimed at detecting and fixing

biases in Spanish texts. (24-CV-3285, ECF 163).[4] Unlike OpenAI, Theodora.AI does not develop

foundational models, pre-train LLMs, or use retrieval augmented generation techniques.

(ECF 271 at 1, 3). Theodora.AI also does not fine tune any of OpenAI's models. (ECF 271-1 ¶ 5).

Though Theodora.AI is broadly positioned in the field of artificial intelligence, the company

conducts its work in a much narrower scope and, thus, is neither an actual nor potential

competitor to OpenAI. OpenAI's apparent view that because Theodora.AI simply exists in the

same space of developing and working on AI-based tools, they must be a competitor with

OpenAI, is wildly overbroad. To hold otherwise would essentially make any private company

---

[4] OpenAI's has also stated that the "majority of [OpenAI's] pretraining data and… alignment data is in English" and "mitigations and measurements [of GPT-4] were mostly designed, built, and tested primarily in English and with a US-centric point of view." (ECF 271 at 1).

even tangentially related to the AI field, as well as any of their C-suite executives, employees, or

vendors, a "competitor" of OpenAI.[5]

Even if Theodora.AI were an actual or potential competitor to OpenAI, OpenAI has failed

to sufficiently allege that Dr. Baeza-Yates is involved in competitive decision making. OpenAI

has provided next to no information as to the extent to which Dr. Baeza-Yates engages in

competitive decision making other than the fact that he serves as CSO of Theodora.AI.

Defendant cites to no case law indicating that serving as CSO in any capacity creates a

presumption of competitive decision making while Plaintiffs indicate that Dr. Baeza-Yates only

works in a very limited advisory capacity for approximately 2 hours a week and does not

participate in the development work. (ECF 271-1). Additionally, the Defendant has not offered

any meaningful reasons to doubt the sufficiency of the existing Protective Order and the

confidentiality provisions therein, particularly when the information is being reviewed by a

company that does not meaningfully compete with them.

Accordingly, the first factor weighs heavily in favor of denying the request for a separate

protective order, and courts generally consider this factor as "arguably the determinative factor

in this analysis." *Infosint S.A. v. H. Lundbeck A.S.*, 06CIV2869LAKRL, 2007 WL 1467784, at *3

(S.D.N.Y. May 16, 2007).

---

[5] *See Advanced Micro Devices, Inc. v. LG Elecs., Inc.*, 14-CV-01012, 2017 WL 3021018, at *3 (N.D. Cal. July 17, 2017) ("However, neither [expert] works directly for an Imagination competitor, and Imagination has provided no concrete evidence that either one engages in or assists with competitive decision-making at an Imagination competitor. The Court finds it unlikely that AMD would be able to select an expert in this space who does not consult, publish, teach, or invent in the relevant field.").

B. Risk of Inadvertent Disclosure

That Theodora.AI is not a competitor to OpenAI greatly informs the Court's analysis for the second and third factor of the *Flores* test because involvement in competitive decision making increases the risk of inadvertent disclosure and, if so, needs to be balanced against the hardship the expert's exclusion would impose." *Infosint S*.A, 2007 WL 1467784, at *3; *see also Errant Gene Therapeutics, LLC v. Sloan-Kettering Institute for Cancer Research*, 15-CV-2044 (AJN) (RLE), 2016 WL 4618972, at *3–4 (S.D.N.Y. Sept. 2, 2016).

Analysis of the risk of inadvertent disclosure is typically triggered only if the first factor is met and "when the proposed recipient of sensitive information has an active relationship with one [of] the disclosing party's competitors." *Flores*, 2021 WL 4441614, at *9. In analyzing risk of inadvertent disclosure, courts have found that protective orders are sufficient to protect against such risk and can counter concerns of competitive injury. *Flores,* 2021 WL 4441614, at *8. Here, because Theodora.AI is not a meaningful competitor with OpenAI, there is little indication that there is a risk of inadvertent disclosure. Even if there were any competition between the two companies, the risk of inadvertent disclosure is minimized given that Dr. Baeza-Yates has signed the existing Protective Order in this case. OpenAI's "*guarantee* of *advertent* disclosure" is conjecture at best, and there is currently no reason to doubt Dr. Baeza-Yates's intent to abide by the Protective Order. Though Defendant argues that it would be "impossible" for Dr. Baeza-Yates to separate the disclosed information from his work at Theodora.AI, fear that an expert could accidentally misuse sensitive information because it is in their head, without more, "is too theoretical to warrant preclusion of that expert's receipt of relevant information." *Flores*, 2021 WL 4441614 at *10.

6

Accordingly, the second factor also weighs in favor of denying the protective order.

C. The Hardship Imposed by the Remedy

The third factor also favors Plaintiffs because granting a protective order will deny them the benefit of having their chosen consultant review the materials with his particular background and expertise.

Plaintiffs will be harmed if Dr. Baeza-Yates, as their chosen consultant, is unable to review the material. (ECF 271 at 3). Defendant argues that Vinay Hooloomann and Professor Oren Etzioni can serve as alternatives to Dr. Baeza-Yates. Their areas of expertise, however, do not entirely overlap with Dr. Baeza-Yates. Mr. Hooloomann is a source code reviewer who does work on the development of LLMs using retrieval-augmented generation (RAG), but his CV indicates that he began such work in April 2023 and works primarily in the context of healthcare. (24-CV-3285, ECF 166 at 194). Dr. Baeza-Yates, on the other hand, has published extensively on web search, web mining, and information retrieval and worked at Yahoo Labs for 10 years doing work on web search, usage data mining, web advertising. (24-CV-3285, ECF 166 at 9–44). Professor Etzioni is primarily involved in research and academia. (24-CV-3285, ECF 166-4, ECF 166-5).[6] Dr. Baeza-Yates's experience on information retrieval outside of academia could provide unique insight and make him "particularly well-suited to address issues related to retrieval-augmented generation, among other issues. . ." (ECF 271 at 3).

Accordingly, the third factor also weighs in favor of denying the protective order.

---

[6] Professor Melanie Mitchell, whose expertise is disclosed in the Nightingale Declaration, was not proffered as an alternative by OpenAI.

D. Timing of the Remedy

The fourth factor favors, albeit slightly, the Defendant. If the remedy of a protective order is granted, it would not substantially impact the efficiency of discovery thus far. If history is any indication of the future of discovery in this case, any delay from the granting of a protective order will be minimal in scope.

E. The Scope of the Remedy

The final factor does not favor either party. The remedy that OpenAI seeks is narrow in that it would only block disclosure to one of Plaintiffs' disclosed experts. However, as discussed above, the requested remedy is also manifestly overbroad because it would effectively render any private company that works in AI a "competitor," and thus block additional experts that Plaintiffs may seek to replace Dr. Baeza-Yates with from accessing material and potentially testifying.

Accordingly, the final factor does not weigh for or against granting a protective order.

F. Weighing the Factors

After considering the five *Flores* factors and having given them the appropriate weight, I find that OpenAI's request for a protective order as to documents produced to Dr. Baeza-Yates is not warranted. Accordingly, Defendant's motion is **DENIED**.

III. CONCLUSION

For the reasons stated above, OpenAI's motion for a protective order to prevent disclosure of confidential information to Dr. Baeza-Yates is **DENIED**.

The Clerk of Court is respectfully directed to close Case No. 24-CV-3285, ECF 163.

SO ORDERED.

Dated: April 1, 2025
      New York, New York

_/s/_ _____

Ona T. Wang
United States Magistrate Judge